JOURNAL OF CLINICAL MICROBIOLOGY, June 2011, p. 2307–2310 Vol. 49, No. 6
0095-1137/11/$12.00 doi:10.1128/JCM.00164-11

Genetic Diversity of *Cryptosporidium* spp. from Bangladeshi Children▽

Kirthi G. Hira,[1]† Melanie R. Mackay,[1]† Andrew D. Hempstead,[2] Sabeena Ahmed,[3] Mohammad M. Karim,[3] Roberta M. O'Connor,[1] Patricia L. Hibberd,[1] Stephen B. Calderwood,[4] Edward T. Ryan,[4] Wasif A. Khan,[3] and Honorine D. Ward[1,2]*

Division of Geographic Medicine and Infectious Diseases, Tufts Medical Center, Boston, Massachusetts[1]; Program in Molecular Microbiology, Sackler School of Graduate Medical Sciences, Tufts University, Boston, Massachusetts[2]; International Center for Diarrhoeal Diseases Research, Dhaka, Bangladesh[3]; and Division of Infectious Diseases, Massachusetts General Hospital, Boston, Massachusetts[4]

Received 26 January 2011/Returned for modification 18 March 2011/Accepted 25 March 2011

The genetic diversity of *Cryptosporidium* spp. from infected children was characterized for the first time in Bangladesh. Seven *C. hominis* and *C. parvum* subtype families (including a new family, IIm) and 15 subtypes (including 2 new subtypes) were identified. The dominance of specific families and subtypes was different from that in other countries.

Cryptosporidium spp. are a major cause of parasitic diarrhea in children under the age of five in developing countries (17, 25). In these countries, cryptosporidiosis in early childhood can lead to persistent diarrhea, growth faltering, and impairment in physical and cognitive development (9, 14). Very little is known about the molecular epidemiology or transmission dynamics of cryptosporidiosis in the developing world, where the burden of cryptosporidiosis is greatest. Two major *Cryptosporidium* species infect humans; *C. hominis* primarily infects humans and is the most frequently identified species in developing countries, and *C. parvum* infects humans as well as animals (29). Recently, subtyping at polymorphic loci has been used to characterize the genetic diversity of *Cryptosporidium* spp. (29). Currently, the most common locus for identifying subtype families and subtypes is that of the *gp40/15* or *gp60* gene, which we and others cloned (8, 22, 26, 27). To date, at least 17 major subtype families, 6 from *C. hominis* (Ia, Ib, Id, Ie, If, Ig) and 11 from *C. parvum* (IIa, IIb, IIc, IId, IIe, IIf, IIh, IIi, IIj, IIk, IIl), have been identified in humans and animals worldwide (18, 29). These include 78 *C. parvum* and 74 *C. hominis* subtypes, classified according to the number and type of serine-coding trinucleotide tandem repeats at the 5′ end of the gene (29).

Although cryptosporidiosis is widely prevalent in Bangladesh (2, 4, 16, 19, 20, 23, 24), there have been no studies on the genetic diversity of *Cryptosporidium* spp. in this country. As part of a case control study on *Cryptosporidium* (identified by screening 1,672 stool samples by microscopy) in children under the age of five presenting with diarrhea to the Dhaka Hospital of the International Centre for Diarrheal Disease Research (ICDDR) in Dhaka, Bangladesh (19), we determined the species and identified the *gp40/15* subtype families and subtypes of *Cryptosporidium* spp. infecting children in the study. The original study (19) was approved by the Ethical Review Committee of the ICDDR, and the use of deidentified stool samples (which were stored at −80°C and shipped to Boston on dry ice) was approved by the Tufts Health Sciences Institutional Review Board.

Stool samples from 46 cases (diarrhea and stool microscopy positive for *Cryptosporidium*) and 46 age-matched controls (diarrhea and microscopy negative for *Cryptosporidium*) at presentation and from 30 cases and 23 controls at follow-up 3 weeks later were analyzed. DNA was extracted from stool samples as described previously (21) or using the QIAamp stool minikit (Qiagen, Inc., Valencia, CA) and analyzed by nested PCR at the 18S rRNA locus (30). Samples from all 46 cases and 7 of 46 controls at presentation and 12 of 30 cases and 2 of 23 controls collected at follow-up 3 weeks later were PCR positive for *Cryptosporidium*. All 67 PCR-positive samples were analyzed by PCR restriction fragment length polymorphism (RFLP) at the 18S rRNA locus for species determination (30). At presentation, 48 of the 53 (91%) samples were positive for *C. hominis*, 4 were positive for *C. parvum* (7%), and 1 was positive for *C. felis* (2%) (Table 1). The same *Cryptosporidium* sp. was identified at the initial and follow-up time points in all 14 of the samples that were PCR positive at follow-up.

To identify subtype families, the *gp40/15* gene was amplified by nested PCR and subjected to PCR RFLP (11). All 48 *C. hominis* samples displayed PCR RFLP profiles corresponding to that of five previously described subtype families (11, 28), including Ia (8 samples), Ib (10 samples), Id (6 samples), Ie (13 samples), and If (11 samples) (Table 1). The *C. felis* sample (number 76) displayed a subtype IIa profile. All 4 *C. parvum* samples (numbers 7, 9, 21, and 78) displayed an RFLP profile that has not been described previously (data not shown). At the follow-up visit, all 12 samples that remained PCR positive for *Cryptosporidium* spp. had the same PCR RFLP profile as at the initial visit except for two that changed from Ib to Id (number 11) and Id to If (number 40) (Table 1), suggesting that these children were reinfected with a different subtype.

To confirm the PCR RFLP findings and identify subtypes,

* Corresponding author. Mailing address: Division of Geographic Medicine and Infectious Diseases, Tufts Medical Center, Box 041, 800 Washington Street, Boston, MA 02111. Phone: (617) 636-7022. Fax: (617) 636-5292. E-mail: hward@tuftsmedicalcenter.org.
† These two authors contributed equally to this work.
▽ Published ahead of print on 6 April 2011.

TABLE 1. *Cryptosporidium* species, subtype families, and subtypes identified from children in the study[d]

Sample no.	Species identified according to 18S rRNA	*gp40/15*	
		Subtype family	Subtype
1	*C. hominis*	Ie	ND
2	*C. hominis*	Id	A15G1
3	*C. hominis*	If	A13G1
4	*C. hominis*	Ib	A8G3
5	*C. hominis*	Ie	A11G3T3
6	*C. hominis*	Ie	ND
7	*C. parvum*	IIm	A7G1
8	*C. hominis*	Ia	A26R2
9	*C. parvum*	IIm	A7G1
10	*C. hominis*	If	A13G1
11	*C. hominis*	Ib[b]	A15G1
12	*C. hominis*	Ie	A11G3T3
13	*C. hominis*	Ib[a]	A9G3
14	*C. hominis*	If	A13G1
15	*C. hominis*	If	A13G1
16	*C. hominis*	Ia	A11R3
17	*C. hominis*	If	A13G1
18	*C. hominis*	If	A13G1
19	*C. hominis*	Ia[c]	A18R3
20	*C. hominis*	Ib	A9G3
21	*C. parvum*	IIm	A7G1
22	*C. hominis*	Id	A15G1
23	*C. hominis*	Ib[a]	A9G3
24	*C. hominis*	Ib[a]	A9G3
25	*C. hominis*	Id	A15G1
26	*C. hominis*	Id	A24
27	*C. hominis*	If	A13G1
28	*C. hominis*	Ie[a]	A11G3T3
29	*C. hominis*	If[a]	A13G1
30	*C. hominis*	Ie	A11G3T3
31	*C. hominis*	Ia	A21R3
32	*C. hominis*	If	A13G1
33	*C. hominis*	If	A13G1
34	*C. hominis*	Ie[a]	A11G3T3
35	*C. hominis*	Ia	A11R3
36	*C. hominis*	Ia	A18R3
37	*C. hominis*	Ie	A11G3T3
38	*C. hominis*	Ib[a]	A7G2
39	*C. hominis*	Ie[a]	A8G1T3
40	*C. hominis*	Id[c]	A16
41	*C. hominis*	Ia[a]	A25R3
42	*C. hominis*	If	A13G1
43	*C. hominis*	Ib	A9G3
44	*C. hominis*	Ia	A21R3
45	*C. hominis*	Ie	A11G3T3
46	*C. hominis*	Ib	A9G3
70	*C. hominis*	Ib	ND
76	*C. felis*	IIa	A17G2R1
77	*C. hominis*	Ie	A11G3T3
78	*C. parvum*	IIm	A7G1
80	*C. hominis*	Ie	A11G3T3
82	*C. hominis*	Id[a]	A16
89	*C. hominis*	Ie	A11G3T3

[a] Same subtype family at initial and follow-up visits.
[b] Subtype family changed from Ib to Id from initial to follow-up visit.
[c] Subtype family changed from Id to If from initial to follow-up visit.
[d] Samples 1 to 46 were microscopy positive (cases), and samples 70, 76, 77, 78, 80, 82, and 89 were microscopy negative (controls) for *Cryptosporidium* spp. at the initial visit. ND, not determined.

the *gp40/15* PCR amplicons from samples obtained at presentation were sequenced (11). In addition, amplicons from the four samples displaying the new PCR RFLP profile were cloned into the pCR 2.1-TOPO vector (Invitrogen Corp., Carlsbad, CA), and the inserts from two or three clones were sequenced. Sequences were obtained from all but three (sample numbers 1, 6, and 70) of the 53 samples (29). Phylogenetic analysis of sequences compared to those deposited in GenBank (Fig. 1) supported the classification of the samples into six previously described *gp40/15* subtype families (Ia, Ib, Id, Ie, If, and IIa). Sequences of the four samples with the new PCR RFLP pattern were identical to each other and were most similar to the subtype family IIe sequence (Fig. 1). However, in addition to five additional tandem TCA repeats in these sequences compared to the IIe sequence, there were several polymorphisms outside the serine-coding region which resulted in amino acid changes (not shown). We therefore propose that these 4 sequences be classified into a new subtype family named "IIm" according to currently accepted nomenclature (29).

Subtype families Ie and If were the most common and were present in 25% and 23% of all samples, respectively, although the overall dominance rates of these subtype families in humans worldwide are only 5.1 and 1.1%, respectively (18). Within subtype family Ie, as reported previously from other areas (18), subtype IeA11G3T3 was dominant. However, all 11 subtypes within the If family were IfA13G1, which has not been reported previously. Subtype families Ib and Ia were each present in 15% of samples in our study. However, the most common subtype worldwide that is present in the Ib subtype family, i.e., IbA10G2 (18), was not identified. As previously reported from other areas (18), subtype family Ia was the most diverse and included five different subtypes. However, again, the most common Ia subtypes, IaA12G1R1 and IaA21G1R1 (18), were not present. Within subtype family Id, subtype IdA15G1 was the most common, as reported previously (18). However, one sample from this subtype family displayed an IdA24 subtype which, again, has not been previously reported. Subtype family Ig was not identified in any of the samples in our study.

All four *C. parvum* samples in our study were of the IIm subtype family. This subtype family has not been identified in any animal thus far and may therefore be another of the so-called "human-adapted" or "anthroponotic" *C. parvum* subtype families, similar to the IIc subtype family (29), which, interestingly, was not identified in our study. Together with the finding that all but one (which was *C. felis*) of the other species identified were *C. hominis*, this suggests that transmission in children in this area is predominantly anthroponotic, i.e., transmitted from human to human, and is similar to that in other developing countries where *C. hominis* and anthroponotic *C. parvum* subtype families predominate (1, 6, 21). Interestingly, although the *gp40/15* locus is not thought to be amplifiable from *C. felis* DNA (29), DNA from the *C. felis* sample from this study and two others from India (1) did amplify with the *gp40/15* primers used in these studies, and all three were of the subtype IIa family. This finding could also be the consequence of mixed infections, as previously described (5).

In conclusion, this is the first study to characterize genetic diversity at the subtype level in *Cryptosporidium* spp. from

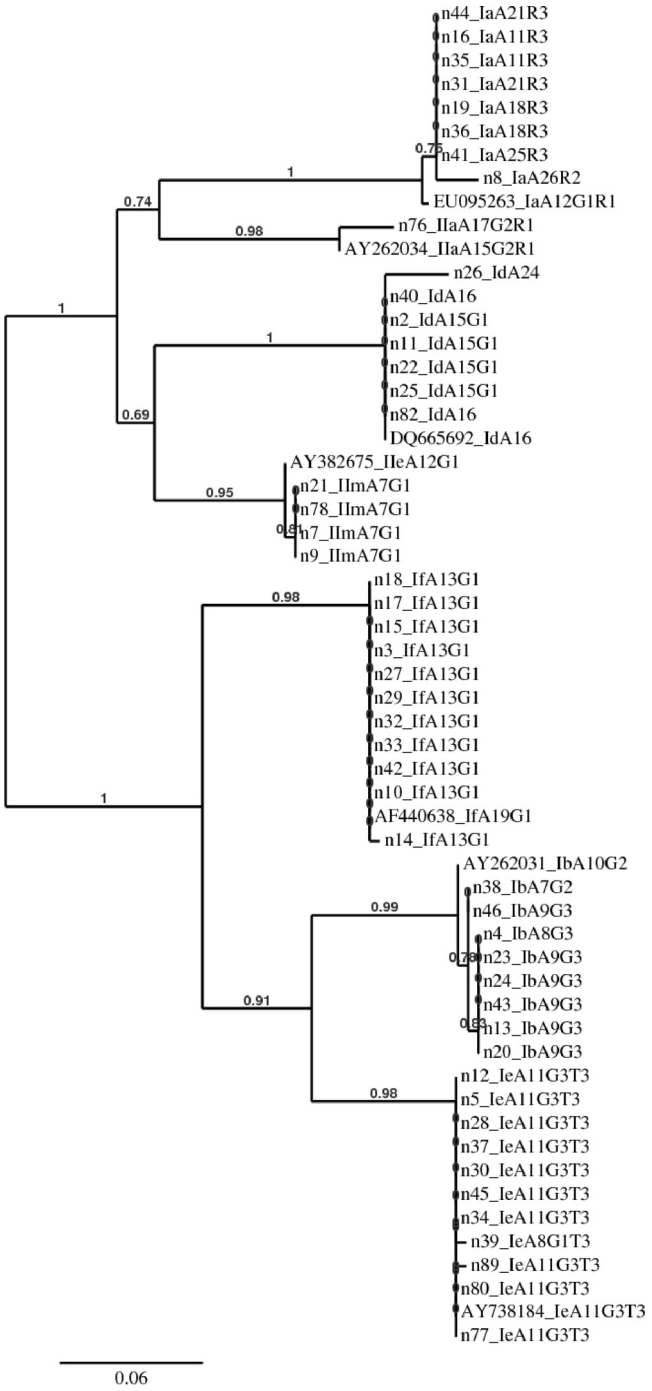

FIG. 1. Phylogenetic analysis of *gp40/15* sequences. All *gp40/15* nucleotide sequences in the study and representative sequences deposited in GenBank were subjected to phylogenetic analysis by the maximum likelihood method using default settings in the Phylogeny.fr server http://www.phylogeny.fr/ (12). These included multiple sequence alignment using MUSCLE (13), alignment curation using G blocks (7), phylogeny using PhyML 3.0 (15), and tree rendering using TreeDyn (10). The numbers at nodes indicate branch support values assessed using the approximate likelihood ratio test (3). Samples from the study are indicated by the letter "n" followed by the sample number. Representative sequences from GenBank are indicated by the accession number. The subtype for each sequence is indicated following the sample number (n) or GenBank accession number.

Bangladesh. Further molecular, clinical, and epidemiological studies of *Cryptosporidium* infections in vulnerable human populations as well as in domestic animals and environmental samples are required to investigate the transmission dynamics of cryptosporidiosis and design effective strategies to block transmission and prevent spread of the disease in developing countries, where the burden of this disease is greatest.

Nucleotide sequence accession numbers. The *gp40/15* sequences obtained in this study have been deposited in GenBank under accession numbers AY700385 to AY700401 and JF927169 to JF927200. Accession numbers of the IIm sequences include AY700401 (number 9), AY700385 (number 21), AY700395 (number 78), and AY700396 (number 7).

This study was supported by an opportunity pool grant and in part by grants UO1 AI45508 (H.D.W.), RO1 AI52786 (H.D.W.), and U01 AI058935 (E.T.R. and S.B.C.), all from the National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH), and K24AT003683 (P.L.H.) from the National Center for Complementary and Alternative Medicine, NIH. K.G.H. and M.R.M. were supported by grant D43TW005571 from the Fogarty International Center, NIH. A.D.H. is supported by grant T32 GM007310 from the National Institute of General Medical Sciences, NIH.

We thank Lihua Xiao, Centers for Disease Control, Atlanta, GA, for helpful suggestions about classification of the IIm subtype family. We thank all the study participants and field and laboratory staff for participating in this study.

REFERENCES

1. **Ajjampur, S. S., et al.** 2007. Molecular and spatial epidemiology of cryptosporidiosis in children in a semiurban community in South India. J. Clin. Microbiol. **45:**915–920.
2. **Albert, M. J., A. S. Faruque, S. M. Faruque, R. B. Sack, and D. Mahalanabis.** 1999. Case-control study of enteropathogens associated with childhood diarrhea in Dhaka, Bangladesh. J. Clin. Microbiol. **37:**3458–3464.
3. **Anisimova, M., and O. Gascuel.** 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Syst. Biol. **55:**539–552.
4. **Bhattacharya, M. K., T. Teka, A. S. Faruque, and G. J. Fuchs.** 1997. Cryptosporidium infection in children in urban Bangladesh. J. Trop. Pediatr. **43:**282–286.
5. **Cama, V., et al.** 2006. Mixed *Cryptosporidium* infections and HIV. Emerg. Infect. Dis. **12:**1025–1028.
6. **Cama, V. A., et al.** 2008. *Cryptosporidium* species and subtypes and clinical manifestations in children, Peru. Emerg. Infect. Dis. **14:**1567–1574.
7. **Castresana, J.** 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. **17:**540–552.
8. **Cevallos, A. M., et al.** 2000. Molecular cloning and expression of a gene encoding *Cryptosporidium parvum* glycoproteins gp40 and gp15. Infect. Immun. **68:**4108–4116.
9. **Checkley, W., et al.** 1998. Effects of *Cryptosporidium parvum* infection in Peruvian children: growth faltering and subsequent catch-up growth. Am. J. Epidemiol. **148:**497–506.
10. **Chevenet, F., C. Brun, A. L. Banuls, B. Jacq, and R. Christen.** 2006. TreeDyn: towards dynamic graphics and annotations for analyses of trees. BMC Bioinformatics **7:**439.
11. **Cohen, S., et al.** 2006. Identification of Cpgp40/15 type Ib as the predominant allele in isolates of *Cryptosporidium* spp. from a waterborne outbreak of gastroenteritis in South Burgundy, France. J. Clin. Microbiol. **44:**589–591.
12. **Dereeper, A., et al.** 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. **36:**W465–W469.
13. **Edgar, R. C.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32:**1792–1797.
14. **Guerrant, D. I., et al.** 1999. Association of early childhood diarrhea and cryptosporidiosis with impaired physical fitness and cognitive function four-seven years later in a poor urban community in northeast Brazil. Am. J. Trop. Med. Hyg. **61:**707–713.
15. **Guindon, S., and O. Gascuel.** 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52:**696–704.
16. **Haque, R., et al.** 2009. Prospective case-control study of the association between common enteric protozoal parasites and diarrhea in Bangladesh. Clin. Infect. Dis. **48:**1191–1197.
17. **Huang, D. B., and A. C. White.** 2006. An updated review on *Cryptosporidium* and *Giardia*. Gastroenterol. Clin. North Am. **35:**291–314, viii.
18. **Jex, A. R., and R. B. Gasser.** 2010. Genetic richness and diversity in *Cryptosporidium hominis* and *C. parvum* reveals major knowledge gaps and a

need for the application of "next generation" technologies—research review. Biotechnol. Adv. **28:**17–26.
19. **Khan, W. A., et al.** 2004. Cryptosporidiosis among Bangladeshi children with diarrhea: a prospective, matched, case-control study of clinical features, epidemiology and systemic antibody responses. Am. J. Trop. Med. Hyg. **71:**412–419.
20. **Kirkpatrick, B. D., et al.** 2008. Association between *Cryptosporidium* infection and human leukocyte antigen class I and class II alleles. J. Infect. Dis. **197:**474–478.
21. **Leav, B. A., et al.** 2002. Analysis of sequence diversity at the highly polymorphic Cpgp40/15 locus among *Cryptosporidium* isolates from human immunodeficiency virus-infected children in South Africa. Infect. Immun. **70:** 3881–3890.
22. **Priest, J. W., J. P. Kwon, M. J. Arrowood, and P. J. Lammie.** 2000. Cloning of the immunodominant 17-kDa antigen from *Cryptosporidium parvum*. Mol. Biochem. Parasitol. **106:**261–271.
23. **Rahman, M., et al.** 1990. Cryptosporidiosis: a cause of diarrhea in Bangladesh. Am. J. Trop. Med. Hyg. **42:**127–130.
24. **Shahid, N. S., A. S. Rahman, and S. C. Sanyal.** 1987. *Cryptosporidium* as a pathogen for diarrhoea in Bangladesh. Trop. Geogr. Med. **39:**265–270.
25. **Snelling, W. J., et al.** 2007. Cryptosporidiosis in developing countries. J. Infect. Dev. Ctries. **1:**242–256.
26. **Strong, W. B., J. Gut, and R. G. Nelson.** 2000. Cloning and sequence analysis of a highly polymorphic *Cryptosporidium parvum* gene encoding a 60-kilodalton glycoprotein and characterization of its 15- and 45-kilodalton zoite surface antigen products. Infect. Immun. **68:**4117–4134.
27. **Winter, G., A. A. Gooley, K. L. Williams, and M. B. Slade.** 2000. Characterization of a major sporozoite surface glycoprotein of *Cryptosporidum parvum*. Funct. Integr. Genomics **1:**207–217.
28. **Wu, Z., I. Nagano, T. Boonmars, T. Nakada, and Y. Takahashi.** 2003. Intraspecies polymorphism of *Cryptosporidium parvum* revealed by PCR-restriction fragment length polymorphism (RFLP) and RFLP-single-strand conformational polymorphism analyses. Appl. Environ. Microbiol. **69:**4720–4726.
29. **Xiao, L.** 2010. Molecular epidemiology of cryptosporidiosis: an update. Exp. Parasitol. **124:**80–89.
30. **Xiao, L., et al.** 1999. Phylogenetic analysis of *Cryptosporidium* parasites based on the small-subunit rRNA gene locus. Appl. Environ. Microbiol. **65:**1578–1583.